Philosophical Magazine,
Vol. 87, Nos. 14–15, 11–21 May 2007, 2191–2220

Discovery, characterization and modelling of novel shape memory behaviour of fcc metal nanowires

W. LIANG and M. ZHOU*

The George W. Woodruff School of Mechanical Engineering,
The School of Materials Science and Engineering,
Georgia Institute of Technology, Atlanta, GA 30332-0405, USA

(*Received 14 September 2006; accepted in revised form 13 February 2007*)

Novel shape memory behaviour was discovered recently in single-crystalline fcc nanowires of Cu, Ni and Au with lateral dimensions below 5 nm. Under proper thermomechanical conditions, these wires can recover elongations up to 50%. This phenomenon only exists at the nanoscale and is associated with reversible lattice reorientations within the fcc lattice structure driven by surface stresses. Whereas the propagation of partial dislocations and twin planes specific to fcc metals are the required mechanism, only materials with higher propensities for twinning (e.g. Cu and Ni) show this behaviour and those with lower propensities for twinning (e.g. Al) do not. This paper provides an overview of this novel behaviour with a focus on the transformation mechanism, driving force, reversible strain, size and temperature effects and energy dissipation. A mechanism-based micromechanical continuum model for the tensile behaviour is developed. This model uses a decomposition of the lattice reorientation process into a reversible, smooth transition between a series of phase-equilibrium states and a superimposed irreversible, dissipative propagation of a twin boundary. The reversible part is associated with strain energy functions with multiple local minima and quantifies the energy conversion process between the twinning phases. The irreversible part is due to the ruggedness of the strain energy landscape, associated with dislocation nucleation, gliding and annihilation, and characterizes the dissipation during the transformation. This model captures all major characteristics of the behaviour, quantifies the size and temperature effects and yields results which are in excellent agreement with data from molecular dynamics simulations.

1. Introduction

Until recently, the shape memory effect (SME) was considered unique to shape memory alloys (SMAs), liquid crystal elastomers and piezoelectric ceramics [1]. Liang and Zhou [2] and Liang *et al.* [3] reported the discovery of a novel shape memory behaviour and its underlying pseudoelasticity in single-crystalline Cu nanowire. Specifically, above a critical temperature T_{cr}, Cu wires exhibit a pseudoelastic behaviour with spontaneously reversible strains up to 50%, well beyond the 5–8% reversible strains typical for most bulk SMAs, as shown in figure 1.

*Corresponding author. Email: min.zhou@gatech.edu

Philosophical Magazine
ISSN 1478–6435 print/ISSN 1478–6443 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/14786430701280943

Below T_{cr}, the deformation is not spontaneously recoverable and the wires retain their deformed configuration after unloading. Subsequent heating to a temperature above T_{cr} activates the SME and allows the wires to return to their original configurations, as shown in figure 2. For Cu wires between 1.3 and 3.2 nm,

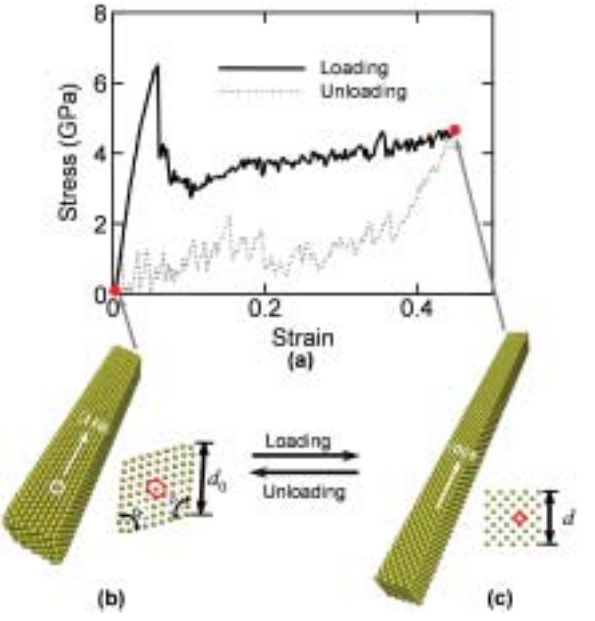

Figure 1. The pseudoelastic behaviour of fcc metal nanowires: (a) the stress–strain hysteresis under tensile loading and unloading; (b) the initial ⟨110⟩/{111} wire configuration; (c) the ⟨001⟩/{001} wire configuration.

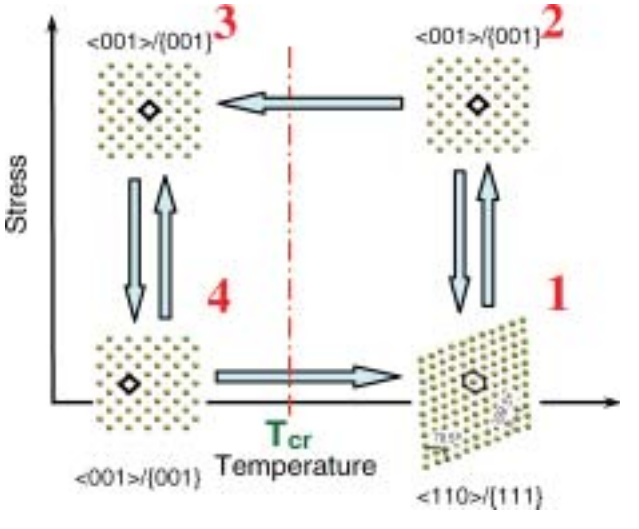

Figure 2. The structural transformation of nanowires under thermomechanical loading; the path for shape memory effect is $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 1$; the path for pseudoelasticity is $1 \rightarrow 2 \rightarrow 1$.

the critical temperature for activating this shape memory is both size- and material-dependent and varies between 50 and 1000 K. This novel behaviour is associated with a reversible lattice reorientation within the fcc crystalline structure. The driving force of this reversible lattice reorientation is the high surface-stress-induced internal compressive stress at the nanoscale. This behaviour only exists at the nanoscale and has not been seen in the corresponding bulk forms of these metals. Following the initial discovery and characterization, Park *et al*. [4], Park and Ji [5] and Liang and Zhou [6] confirmed a similar shape memory behavior in Ni and Ag wires with the same transformation mechanism.

Not all fcc metal nanowires show this shape memory. For example, the tensile deformation of Al nanowires is irreversible upon unloading, regardless of temperature [6]. On the other hand, Au wires exhibit transitional behaviours from pseudoelasticity like that of Cu and Ni wires at low temperatures to irreversible deformation like that of Al wires at high temperatures [4, 6]. Liang and Zhou pointed out that whether a metal nanowire shows the shape memory is determined by the competition between twinning and slip, two related but different deformation mechanisms [6]. Specifically, the pseudoelasticity of Cu and Ni wires at the nanometre scale is associated with a reversible lattice reorientation process through twin boundary propagation. On the other hand, the irreversible tensile deformation of Al wires is due to crystalline slip. The competition between twinning and slip is important in understanding the pseudoelastic behaviour and the SME of the fcc nanowires analyzed here. Twinning leads to pseudoelasticity and slip leads to permanent deformation. Liang and Zhou [6] found that the difference in the deformation mechanisms can be well explained by the twinnability parameter identified by Tadmor and Hai which measures the tendency of a material for twinning [7]. Specifically, fcc metals with higher twinnability values (such as Cu, Ni, and Au) favour twinning through $a/6\langle\bar{2}11\rangle$-type partial dislocations which leads to the pseudoelastic behaviour and SME in their nanowires. On the other hand, fcc metals with low twinnability values (such as Al) favour slip through $a/2\langle\bar{1}10\rangle$-type full dislocations which leads to irreversible plastic deformations even at the nanoscale. Au has a transitional twinnability value which is lower than those of Cu and Ni but higher than Al. Therefore, Au nanowires show mixed deformation mechanisms which are temperature-dependent.

An interesting question is why the novel SME only exists in nanowires, but not in bulk fcc metals. Liang and Zhou showed that a combination of three factors is at work [6]. These factors are: (i) only at the nanoscale is the surface-stress-induced internal compressive stress high enough (on the order of $3 \sim 8$ GPa) to initiate a spontaneous lattice reorientation without external influence; (ii) the reversibility of the lattice reorientation or twin boundary propagation process results from the unique slender 1D structure of the nanowires; (iii) twinning becomes a viable deformation mode only at the nanoscale, in nanocrystalline metals and nanowires. It is well established that slip is the favoured deformation mechanism in bulk fcc metals at ambient temperatures. For nanowires however, the orientation of lateral surfaces has a direct and first order effect on the operant deformation mode in nanowires, as pointed out by Park *et al*. [8]. This is especially the case for fcc metals with relatively higher tendencies for twinning as quantified by, e.g. the twinnability parameter which by itself does not account for size-induced effects at the nanoscale from

features such as surfaces and grain boundaries. Another way to look at the issue is, as will be shown later in this paper, the size-dependence of deformation mechanism has to do with the equilibrium width of stable stacking faults which is on the order of nanometres for most fcc metals. Not coincidentally, the novel SME and the underlying pseudoelasticity are observed only when the characteristic lengths (lateral dimensions) of the nanowires are on the same order of the size of the stacking faults.

This paper focuses on the discovery, characterization and modelling of the novel shape memory behaviour in fcc metal nanowires. In section 2, the structures and structural stabilities of fcc nanowires are discussed within the framework of a top-down fabrication process. In section 3, the pseudoelastic behaviour of nanowires is characterized with emphasis on the transformation mechanism, energy dissipation and the quantification of large transformation strains. In section 4, we explain why not all fcc metal nanowires show the shape memory by analyzing the competition between twinning and slip. In section 5, the shape memory effect in nanowires is characterized with a focus on surface effect and driving force for lattice reorientation. Based on an analysis of deformation mechanisms, a micromechanical continuum model is developed in section 6 to characterize the unique lattice reorientation process. Finally, in section 7 a summary of the current key findings is provided.

2. Wire structures

The shape memory nanowires have a unique structure with ⟨110⟩ axes and {111} lateral surfaces (hereafter denoted as the ⟨110⟩/{111} structure or configuration). This structure has been frequently observed in experiments [9–13] and in atomistic simulations [14–17]. The preference of this wire configuration is due to the significant surface effects in nanowires. Specifically, the surface energy of nanowires constitutes a significant portion of the total free energy because of their extremely high surface-to-volume ratios compared to bulk materials. For example, the surface-to-volume ratio of a nanowire with a diameter of 4 nm is 10^6 times that of the typical macroscopic tensile specimen with a diameter of 4 mm. Consequently, crystalline nanowires tend to assume axes and surfaces in preferred orientations to lower their total energy. It has been observed that most fcc metal wires have {111} and {001} surfaces and axes in ⟨111⟩, ⟨110⟩, or ⟨001⟩ directions [18, 19]. Due to the strong orientation preferences, lattice reconstruction frequently occurs when initial configurations do not conform to the preferred conditions. Under such circumstances, surfaces with higher energies reorganize into surfaces with lower energies. Examples include reconstructions of {110} surfaces into {001} surfaces [9] and {001} surfaces into {111} surfaces [20]. Computationally, stable wire structures can be obtained by simulating the "top-down" fabrication process of wires, as shown in figure 3a. Specifically, wires with ⟨001⟩ axes and {001} surfaces (hereafter called the ⟨001⟩/{001} structure or configuration) are isolated from perfect bulk lattices and are allowed to relax at a constant temperature. During relaxation of a wire, lattice reconstruction occurs and {001} lateral surfaces reorganize into lower-energy

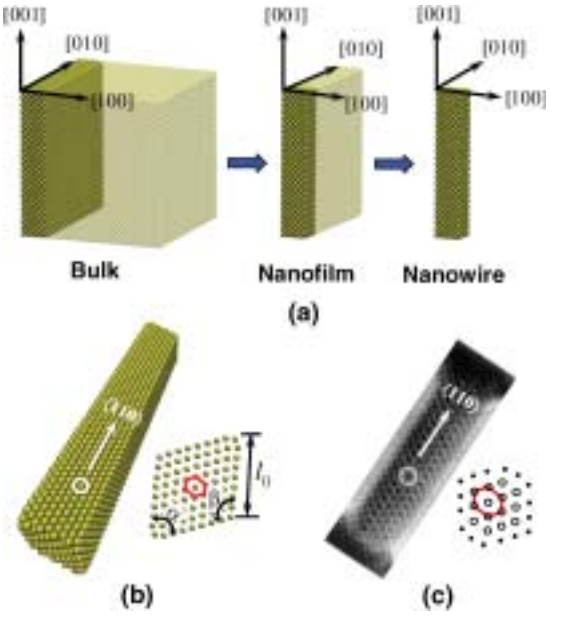

Figure 3. (a) A schematic illustration of the top-down fabrication of nanowires; (b) a stable ⟨110⟩/{111} Au wire predicted by MD simulations; (c) a ⟨110⟩/{111} Au wire cut from a {001} nanofilm by electron beam irradiation [20].

{111} surfaces. Meanwhile, the ⟨001⟩ axis evolves into a ⟨110⟩ axis, resulting in a free standing nanowire with the ⟨110⟩/{111} configuration. This unique wire configuration has been consistently observed in experimentally fabricated wires (through both top-down and bottom-up approaches) of Cu, Au, and Pt. Examples include Cu wires fabricated via vacuum vapour deposition [11] and by a surfactant-assisted hydrothermal reduction process [12] and Au wires fabricated by cutting Au nanofilms via electron beam eradiation [20]. Just like what is seen in simulations, the laboratory nanowires have defect-free single-crystalline fcc structures with ⟨110⟩ axes and {111} surfaces, indicating that ⟨110⟩ is the preferred growth orientation and {111} planes are preferred lateral surfaces, as shown in figures 3b and c.

3. Pseudoelasticity

3.1. *Stress–strain curves and lattice reorientation*

Upon tensile loading and unloading, ⟨110⟩/{111} wires exhibit a novel pseudoelastic behaviour with reversible strains up to 50%. The large reversible deformation is due to a lattice reorientation process which assigns ⟨110⟩/{111} wires a stress-strain relation drastically different from those of the corresponding bulk metals, as shown

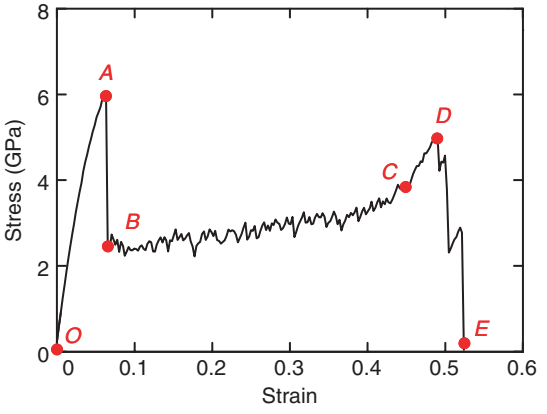

Figure 4. Stress–strain curve of a 1.96×1.96 nm $\langle 110 \rangle / \{111\}$ Cu nanowire during tensile deformation at 300 K.

in figure 4. The stress–strain curve consists of four deformation stages, each with its own underling deformation mechanisms. Specifically,

(1) between O and A, the $\langle 110 \rangle / \{111\}$ wire undergoes elastic stretching;
(2) point A marks the beginning of the lattice reorientation process which leads to a $\langle 001 \rangle / \{001\}$ configuration at point C. This configuration is the same as the original (unrelaxed) $\langle 001 \rangle / \{001\}$ configuration in figures 1c and 3a;
(3) between C and D, the newly formed $\langle 001 \rangle / \{001\}$ wire undergoes elastic stretching;
(4) further loading beyond D causes the wire to yield through the formation of full dislocations which ultimately lead to necking and fracture of the nanowire at E [21].

The unique lattice reorientation process ($A \rightarrow C$ in figure 4) is the key to the pseudoelastic behaviour and the SME of the wires because it is reversible upon unloading. It proceeds through the reversible propagation of a coherent twin boundary separating the $\langle 110 \rangle / \{111\}$ phase and the $\langle 001 \rangle / \{001\}$ phase, as shown in figure 5a. The stacking sequence within each region is ABCABC in the direction perpendicular to the boundary. At the interface, the stacking sequence is ABC|A|CBA where the middle A is the mirror plane. Crystallographically, this is a coherent $\{111\}$ twin plane with its misorientation axis aligned in the $[\bar{1}1\bar{2}]$ direction, as shown in figure 5b. The twin boundary propagates in a "stick-slip" manner through successive dislocation nucleation, gliding, and annihilation. As the twin boundary sweeps through the wire length, the wire is progressively transformed into the $\langle 001 \rangle / \{001\}$ phase. Upon the arrival of the twin boundary at the other end of the wire (corresponding to point C in figure 4), the whole wire is in the $\langle 001 \rangle / \{001\}$ phase [6, 22]. A section of the transformed (lower) region of the wire shows a relative shift of atoms on (001) planes. The magnitude of the shift is approximately one half of the lattice spacing. This shift is part of the overall elastic bending of the wire that accommodates the orientation difference between the lattices on the two sides of the twin boundary and the overall loading direction. This relative shift disappears

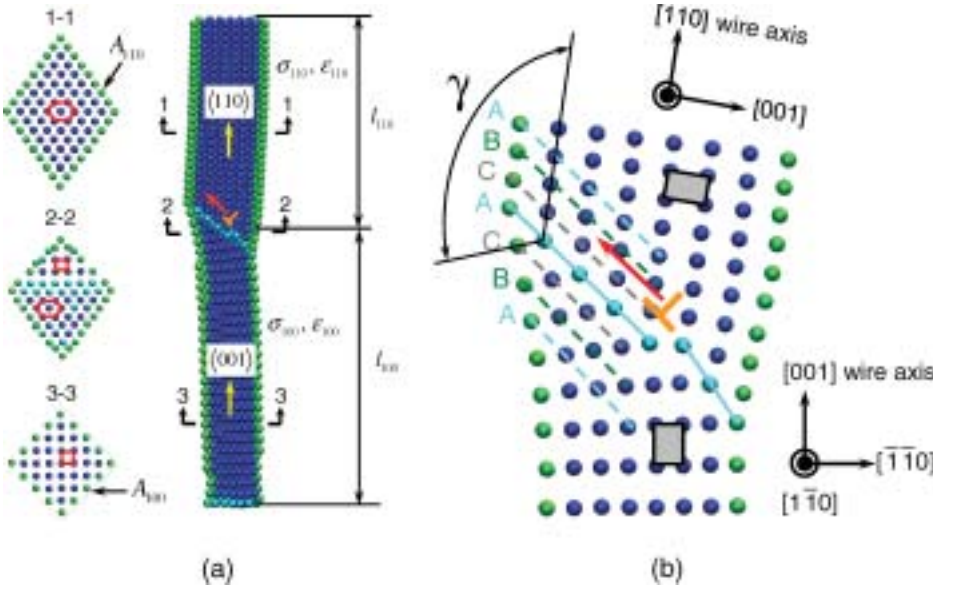

Figure 5. (a) Lattice orientation as viewed on a cross-section of a 1.96×1.96 nm Cu wire at a strain of 0.24; the right hand image shows a sectional view along the wire axis and the $[\bar{1}\bar{1}0]$ diagonal of a cross-section; cross-section 1-1 shows the elongated hexagonal lattice in the unrotated domain within the $\langle 110\rangle/\{111\}$ phase; cross-section 2-2 is in the transitional region containing both the $\langle 001\rangle/\{001\}$ and the $\langle 110\rangle/\{111\}$ phases; and cross-section 3-3 shows the square lattice in the reoriented domain within the $\langle 001\rangle/\{001\}$ phase. (b) A close-up view of the $\frac{1}{6}\langle 112\rangle$-type Shockley partial dislocation and twin interface in (a).

as the loading continues and does not a significantly affect the overall deformation of the wire. If unloading occurs at temperatures above a critical temperature T_{cr} (size-dependent, see Liang and Zhou [16] for specific values), the $\langle 001\rangle/\{001\}$ wire spontaneously transforms back to the $\langle 110\rangle/\{111\}$ configuration. The two-way lattice reorientations between $\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ allow the associated deformations to be fully recovered, giving rise to a pseudoelasticity at the nanoscale.

This novel pseudoelastic behaviour has been observed under dynamic, quasistatic and static loading conditions, as shown in figure 6. The fully dynamic (strain rate $= 10^9\,s^{-1}$) and quasistatic loading conditions are obtained through molecular dynamics simulations and the static loading condition is achieved within the framework of molecular statics. In all three cases, one end of the wire is fixed and a tensile load is applied at the other end. Specifically, for the fully dynamic loading a constant boundary velocity is applied at the moving end. A velocity gradient along the wire axis is specified initially to avoid the generation of a shock wave in the wire. To approximate quasistatic loading, each load step involves all atoms being first displaced according to a prescribed uniform strain increment in the length direction. The wires are then equilibrated at constant temperature using a Nosé–Hoover thermostat [23, 24] with their ends fixed to obtain a macroscopic equilibrium configuration at the prescribed strain [25]. This relaxation process allows structural changes to occur, if the conditions so dictate. Hence, the deformation proceeds

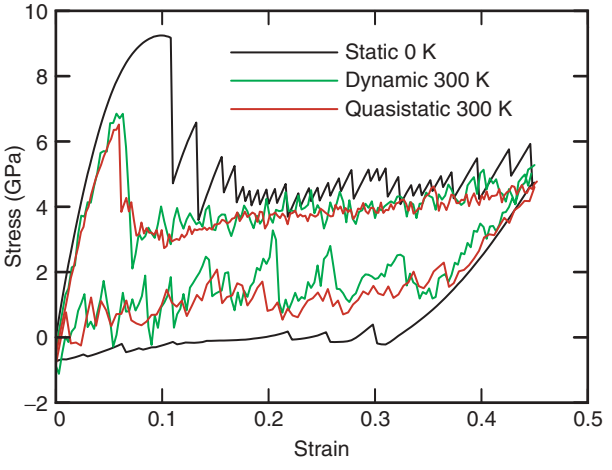

Figure 6. The pseudoelastic behaviour of a 1.52×1.52 nm $\langle 110 \rangle / \{111\}$ Cu nanowire under dynamic ($\dot{\varepsilon} = 1.0 \times 10^9\ s^{-1}$), quasistatic, static loading conditions.

through a series of equilibrium states, approximating a quasistatic deformation process. To effect static loading, a procedure similar to that in the quasistatic case is used, except that the equilibrium state in each loading step is obtained via energy minimization using molecular statics. The static loading is at 0 K while the dynamic and quasistatic loading can be applied at finite temperatures. Figure 6 illustrates that the mechanical behaviours under all three loading conditions show similar characteristics. Indeed, the same deformation mechanisms are at work in all three cases. In particular, the stress-strain relations under dynamic and quasistatic loading (300 K) essentially coincide. On the other hand, the stress level under static loading is higher while the stress level during unloading is lower than the corresponding levels for the dynamic and quasistatic conditions, with the size of the hysteresis loop changing accordingly, indicating more dissipation per cycle at lower temperatures.

3.2. *Energy dissipation*

Heat dissipation is an important factor in many applications, especially under conditions of high state-switching rates. As shown in figures 6 and 7, the responses of the nanowires are associated with relatively large hysteresis loops, indicating a substantial amount of heat is generated in each cycle of deformation. This energy dissipation is primarily due to the nucleation and annihilation of dislocations associated with the two-way propagation of twin boundaries. The amount of dissipation is both temperature- and size-dependent. For the same wire, more energy is dissipated at lower temperatures because the stress is higher during loading and lower during unloading (see the larger hysteresis loop for 0 K in figure 6). Specifically, the energy dissipations are 2.12, 1.15, 1.06 and 1.00 $J\,m^{-3}$ at 0, 100, 200 and 300 K, respectively. The dissipation per unit volume of material is higher at smaller wire sizes, as indicated by the hysteresis loops in figure 8. Specifically, the

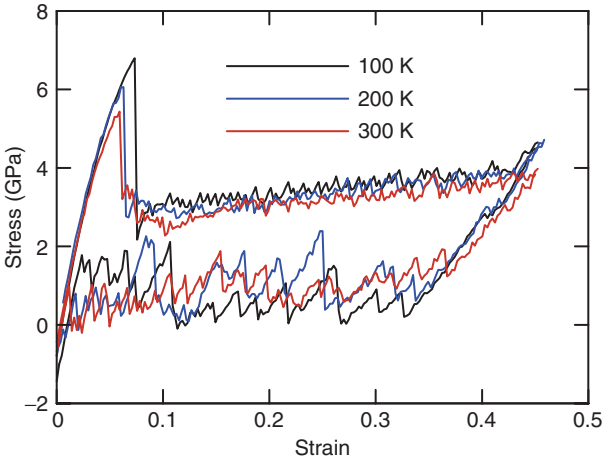

Figure 7. The pseudoelastic behaviour of a 1.52×1.52 nm $\langle 110 \rangle / \{111\}$ Cu wire at 100, 200 and 300 K.

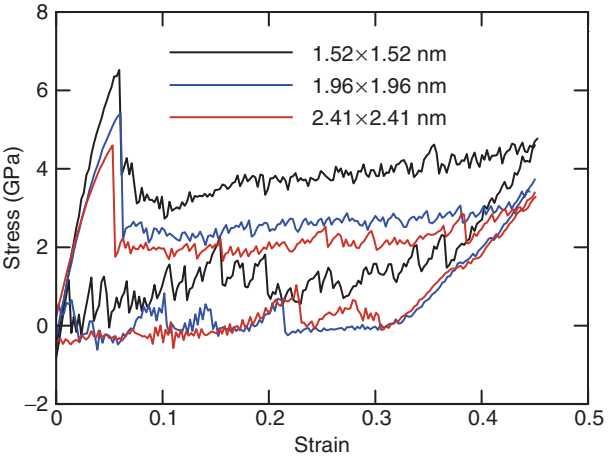

Figure 8. The pseudoelastic behaviours of three $\langle 110 \rangle / \{111\}$ Cu wires of different sizes at 300 K; the lateral dimensions of the three wires are 1.52×1.52 nm, 1.96×1.96 nm, and 2.41×2.41 nm.

dissipation at 300 K is 1.0, 0.96 and 0.81 J m^{-3} for wires with sizes of 1.52×1.52 nm, 1.96×1.96 nm and 2.41×2.41 nm, respectively. One factor contributing to this size dependence of energy dissipation on a per-unit-volume basis is the fact that wires of different sizes have the same number of dislocations nucleated and annihilated and therefore essentially the same total amount of dissipation associated with dislocation cycling.

To put the issue of energy dissipation in perspective, temperature increases under adiabatic conditions with strain rates of 10^8 and 10^9 s^{-1} are analyzed, as shown in figure 9. In the early stage of loading (A $\rightarrow$ B), temperature decreases slightly as

thermal energy is converted into potential energy during the elastic stretching of the ⟨110⟩/{111} wire. The sudden increase in temperature between B and C is associated with the initiation of the lattice reorientation process or the emission of a partial dislocation as a significant amount of energy is dissipated. Subsequent propagation of the twin boundary causes a steady increase in temperature (B → D) as successive dislocation nucleation and annihilation occur. The temperature remains essentially constant between D and E as the ⟨001⟩/{001} wire is unloaded elastically. Between E and F, temperature increases steadily as unloading and relaxation progress through the propagation of multiple twin boundaries (not shown). For the adiabatic conditions analyzed, the temperature increase per load–unload cycle is approximately 150–200 K. Such increases can lead to significant heating during rapid cyclic loading. Obviously, it is very important to effectively channel the heat out of the wires under such conditions. Actual temperature increases should be lower than the adiabatic prediction here due to thermal convection, conduction, and radiation. The high surface-to-volume ratios of the nanowires compared with bulk materials offer a benefit in this regard by facilitating heat transfer out of the wires. Furthermore, many applications may not require the wires to be stretched to

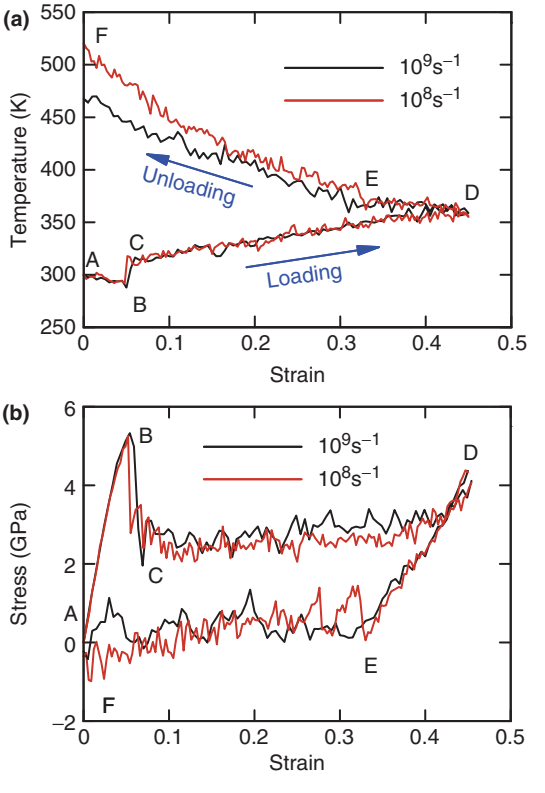

Figure 9. The thermomechanical behaviour of a 1.52 × 1.52 nm Cu wire under adiabatic loading conditions; (a) the temperature increases at two different rates of deformation, and (b) the stress–strain hysteresis loops at two different rates of deformation.

their limit reversible strains, therefore, resulting in less heat generation and lower temperature increases.

3.3. *Large transformation strain*

The forward and reverse lattice reorientations during loading and unloading induce large transformation strains (ε_{tr}) which leads to extraordinarily high levels of reversible strains of up to 50%. ε_{tr} can be quantified by a simple crystallographic analysis, as illustrated in Figure 10 which compares the $(1\bar{1}0)$ planes in both the $\langle 110\rangle/\{111\}$ (point A in figure 4) and the $\langle 001\rangle/\{001\}$ (point C in figure 4) configurations. Clearly, the forward (loading) and backward (unloading) lattice reorientations manifest as 90° rotations in opposite directions of the unit cell in the $(1\bar{1}0)$ plane. The length and width of the rectangular unit cell in both cases are, respectively, a and $(\sqrt{2}/2)a$; where a is the lattice constant in the stressed states and is assumed to be the same at A and C. Hence, the axial strain associated with the lattice reorientation between A and C is

$$\varepsilon_{tr} = \frac{(a - (\sqrt{2}/2)a)}{(\sqrt{2}/2)a} = 0.414. \tag{1}$$

ε_{tr} is an attribute of the fcc structure and is independent of a. Consequently, the pseudoelastic strain associated with the lattice reorientation which constitutes the primary part of the total recoverable strain (ε_r) is the same for wires of all FCC metals and of all sizes. In addition to ε_{tr}, ε_r also includes the elastic strain $\varepsilon^e_{\langle 110\rangle}$ associated with the lattice stretching in the $\langle 110\rangle/\{111\}$ configuration between O and A (figure 4) and the elastic strain $\varepsilon^e_{\langle 001\rangle}$ associated with the lattice stretching in the $\langle 001\rangle/\{001\}$ configuration between C and D, i.e.

$$\varepsilon_r \approx \varepsilon^e_{\langle 110\rangle} + \varepsilon_{tr} + \varepsilon^e_{\langle 001\rangle}. \tag{2}$$

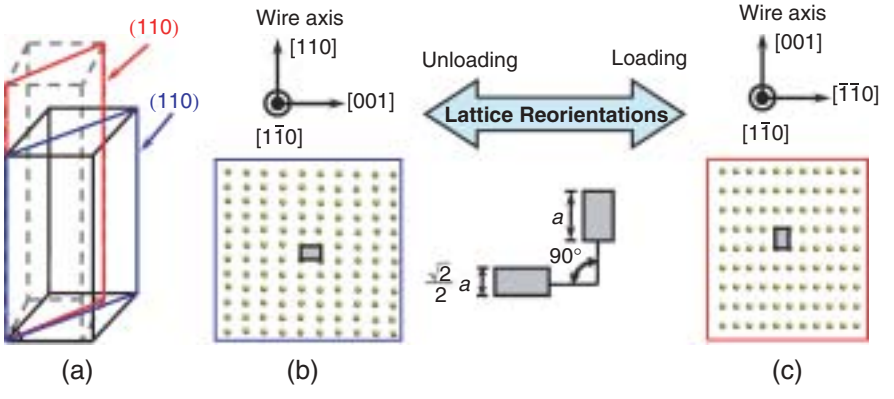

Figure 10. Reversible lattice reorientations upon loading and unloading in metal nanowires: (a) a schematic illustration of the wire configurations before and after lattice reorientation; (b) the $(1\bar{1}0)$ atomic plane highlighted in blue in (a); (c) the same $(1\bar{1}0)$ atomic plane highlighted in red in (a).

Overall, ε_r is around 50% for Cu. Note that, whereas ε_{tr} is material-independent, $\varepsilon^e_{\langle 110\rangle}$ and $\varepsilon^e_{\langle 001\rangle}$ are material-specific. Also, only partial lattice reorientations (only a part of a wire converts from the $\langle 110\rangle/\{111\}$ configuration into the $\langle 001\rangle/\{001\}$ configuration) may occur in some materials (e.g. Au) or at higher temperatures. Therefore, the actual value of recoverable strain may be smaller than 50%, depending on material, wire size, and temperature.

4. Twinning vs. slip

Not all fcc metal nanowires show pseudoelastic behaviours. In particular, it has been found that Al wires do not deform via lattice reorientation like Cu and Ni wires. Instead, they develop a neck which leads to thinning and eventual rupture of the wire. Au nanowires, on the other hand, show a temperature-dependent transitional behaviour between the pseudoelasticity of Cu and Ni wires and the plasticity of Al wires. Specifically, at temperatures below 200 K Au wires show similar forward and reverse lattice reorientations through the propagation of twin boundaries as Cu and Ni wires do. However, the reorientation process during loading may not sweep through the entire wire length and necking can occur at the twin boundary before the reorientation process is complete. At temperatures above 200 K, no reorientation is observed and necking starts immediately after yielding. This scenario corresponds to the case of Al wires discussed above, with the deformation progressing via slip and being irreversible.

This difference in behaviour arises from different deformation mechanisms. While twinning is responsible for the pseudoelastic behaviour in Cu and Ni wires, slip via full dislocations of the $(1/2)\langle 110\rangle$ type is primarily responsible for the necking process in Al wires. The mixed occurrence of twinning and slip explains the transitional behaviour of Au wires between pseudoelasticity and plasticity. Being two competitive mechanisms in fcc metals, twinning and slip are known to occur under different conditions. Conventional wisdom suggests that metals with low stable stacking fault energy γ_{sf} are more likely to deform through twinning. However, γ_{sf} alone is not enough to determine whether a metal would deform via twinning or slip. For example, Ni wires deform via twinning while Al wires deform via slip, even though Ni and Al have similar levels of γ_{sf} (table 1). Swygenhoven *et al.* have shown that the competition between twinning and slip is primarily determined by the energy ratio of γ_{sf}/γ_{us} (γ_{us} is the unstable stacking fault energy) [32]. Specifically, metals with lower γ_{sf}/γ_{us} values are more likely to deform via twinning than slip. This understanding is consistent with the observation that twinning is more likely to occur in Ni wires than in Al wires since $(\gamma_{sf}/\gamma_{us})_{Ni} = 0.317$ and $(\gamma_{sf}/\gamma_{us})_{Al} = 0.836$ (see figure 11). Even though γ_{sf}/γ_{us} is a better parameter than γ_{sf} for assessing the competition between twinning and slip, it cannot explain all cases and a higher order effect appears to exit. For example, as previously shown, twinning is more likely in Ni wires than in Au wires despite the fact that Ni has a slightly higher ratio of γ_{sf}/γ_{us}(0.317) than Au (0.307) (see table 1). The reason is that the competition is also affected by the energy barriers associated with both slip and twinning. γ_{sf}/γ_{us} only

Table 1. A comparison of stacking fault energies.

	γ_{sf} (mJ m^{-2})		EAM						
	EAM	Experiments	γ_{us} (mJ m^{-2})	γ_{ut} (mJ m^{-2})	γ_{sf}/γ_{us}	γ_{us}/γ_{ut}	Twinnability τ_a	d_{sep} (nm)	Behaviour
Cu	45	45[a]	180	202	0.249	0.896	1.040	3.2	PE* & SME
Ni	125	125[b]	402	460	0.317	0.878	1.020	0.56	PE
Au	31	32[c]	101	122	0.307	0.828	0.991	8	Transitional
Al	146	120–166 [d,e]	189	240	0.836	0.788	0.896	0.24	Irreversible

[a]Carter and Ray [26]; [b]Balluffi [27]; [c]Stobbs and Sworn [28]; [d]Murr [29]; [e]Westmaco and Peck [30] and Rautioaho [31]; PE* = pseudoelasticity.

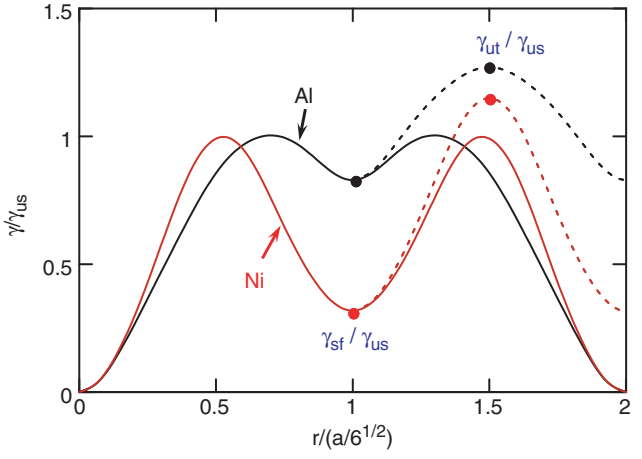

Figure 11. The generalized stacking fault energy curves for Ni and Al normalized by the respective unstable stacking fault energy, γ_{us}.

quantifies the energy barrier for slip and the energy barrier for twinning must also be considered.

The competition between twinning and slip can be better explained by the dislocation nucleation criterion of Rice [33] and the twin nucleation criterion of Tadmor and Hai [7]. In Rice's theory, a $(1/2)\langle 110\rangle$-type full dislocation can be considered as being formed by a leading and a trailing $(1/6)\langle 112\rangle$-type Shockley partial dislocations in two successive steps, i.e.

$$\frac{\mathbf{a}}{6}[\bar{2}11] + \frac{\mathbf{a}}{6}[\bar{1}2\bar{1}] \rightarrow \frac{\mathbf{a}}{2}[\bar{1}10]. \tag{3}$$

Based on Rice's theory, Tadmor and Hai showed that in the twinning mode, a leading partial is first nucleated from a sharp edge of a surface, the same as the first step in the nucleation of dissociated full dislocations [7]. Instead of the emission of a trailing partial, a second leading partial (twinning partial) is nucleated on a neighbouring slip plane next to the original pair of slip planes. Clearly, the

deformation mode (twinning or full dislocation motion) is determined by the competition between the nucleation of the trailing partial needed to complete the dissociated full dislocation and the nucleation of the twinning partial needed to form a microtwin. Based on the aforementioned dislocation nucleation process, Tadmor and Hai [7] developed a criterion for the onset of deformation twinning which quantifies the competition between slip and twinning. The related twinnability is

$$\tau_a = \left[1.136 - 0.151\frac{\gamma_{sf}}{\gamma_{us}}\right]\sqrt{\frac{\gamma_{us}}{\gamma_{ut}}}. \tag{4}$$

This parameter depends not only on γ_{sf} and γ_{us}, but also on γ_{ut} which is the unstable twinning energy [34]. γ_{sf}/γ_{us} and γ_{us}/γ_{ut} measure the energy barriers associated twinning and slip, respectively. On one hand, twinning is favoured when γ_{sf}/γ_{us} is small and γ_{us}/γ_{ut} is large (as is the case for Cu and Ni) because small values of γ_{sf}/γ_{us} indicate higher barriers of $\gamma_{us} - \gamma_{sf}$ for the nucleation of the trailing partial and large γ_{us}/γ_{ut} values indicate lower barriers for the nucleation of a twinning partial. On the other hand, slip is favoured when γ_{sf}/γ_{us} is large and γ_{us}/γ_{ut} is small (as is the case for Al, see figure 11). As shown in table 1, the twinnability ranking based on interatomic potentials for the metals analyzed is

$$\tau_a^{Cu} > \tau_a^{Ni} > \tau_a^{Au} > \tau_a^{Al}. \tag{5}$$

This is in good agreement with the results of experiments and first principle calculations [34]. Furthermore, it clearly explains the different behaviours of the fcc nanowires. Cu and Ni exhibit reversible deformations through twinning because it is favoured over slip. On the other hand, the deformation of Al wires is irreversible because slip is the favoured over twinning.

The above analysis does not account for the effect of temperature which is important for Au because it has a twinnability in the transition regime. Our observation of Au wires showing twinning (and therefore, pseudoelasticity and SME) at lower temperatures and slip (and therefore, plasticity) at higher temperatures is consistent with the experimental observation that twinning tends to occur at lower temperatures (see, for example, Hertzberg [35] and Meyers [36]).

Whereas twinning is vital to the pseudoelasticity in the nanowires, slip is generally favoured over twinning in bulk fcc metals under normal loading conditions. Twinning becomes a viable deformation mode only when the dominant characteristic length (e.g. grain size or wire size) scale approaches nanometers. Both experiments and simulations have shown that twinning is the primary deformation mechanism in nanocrystalline metals, nanowires and nanoparticles [37]. Although τ_a characterizes the tendency for twinning, it does not account for the effect of size on deformation mechanisms. Park *et al.* pointed out that at the nanometre scale the crystallographic nature of side surfaces can have a direct, first order effect on the operant deformation mode [8]. The deformation mechanism which tends to generate surfaces with lower energies is favoured. It is important to note that the ability to completely alter the operant deformation mechanisms in a material via rotation about its uniaxial deformation direction is a unique effect at the nanometre scale that does not occur in bulk solids.

In addition to side surfaces, the size dependence of deformation mechanisms is also related to the equilibrium width of stable stacking faults which are bound by Shockley partial dislocations formed through the dissociation of full $(1/2)\langle 110\rangle$ dislocations, as described by equation (3). The separation between the two partials (d_{sep}) or the width of a stacking fault is primarily determined by the balance of elastic repulsion between the partials and attraction due to the interruption of the perfect fcc stacking sequence. This separation between the two partials corresponding to a dissociated perfect edge dislocation is

$$d_{\text{sep}} = \frac{Gb_1b_2}{8\pi\gamma_{\text{sf}}}\frac{2+v}{1-v}, \tag{6}$$

where G is the shear modulus, v is the Poisson's ratio, b_1 and b_2 are the magnitudes of the Burger's vectors of the two Shockley partials, respectively [38]. Here, the two partials are considered separate dislocations connected by the intrinsic stacking fault between them. Alternatively, the whole assembly of the two partials and the stacking fault can be viewed as the extended core of a full dislocation with a core size of d_{sep} which is inversely proportional to γ_{sf}. Obviously, d_{sep} is smaller in metals with higher γ_{sf} and larger in metals with lower γ_{sf}. Overall, both d_{sep} and the wire sizes of interest are on the order of nanometres (table 1). When d_{sep} is larger than the lateral dimensions of a nanowire, the cores of full dislocations cannot exist within the wire, hence slip is unlikely to occur. Consequently, twinning becomes more likely in nanocrystalline metals or nanowires. On the other hand, d_{sep} is far smaller than the characteristic lengths (usually the grain sizes or macroscopic dimensions) of bulk metals and consequently has little effect on deformation mechanism, allowing slip to be favoured.

5. Shape memory

Like the martensitic transformation in normal bulk SMAs, the spontaneous lattice reorientation is strongly temperature-dependent. Specifically, the reverse lattice reorientation from $\langle 001\rangle$ to $\langle 110\rangle$ occurs only above a critical temperature T_{cr} (which is size-dependent). If unloading takes place at temperatures below T_{cr}, the reverse lattice reorientation does not occur and the wire retains the $\langle 001\rangle/\{001\}$ configuration. When subsequently heated above T_{cr}, the unloaded $\langle 001\rangle/\{001\}$ wire spontaneously returns to its original $\langle 110\rangle/\{111\}$ configuration through the reverse lattice reorientation. This is a one-way SME with the $\langle 110\rangle/\{111\}$ configuration being the parent state described in figure 2.

5.1. *Surface effect and driving force*

The driving force for the martensitic transformation in bulk SMAs is the difference in the free energies of the parent and product phases. What is the driving force for the spontaneous lattice reorientation, since both $\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ configurations have the fcc crystalline structure? The answer lies in the surfaces

and the extremely high surface-to-volume ratios of nanowires which can significantly affect structural stability. Surface energy constitutes a significant portion of the total strain energy of the wire. {111} surfaces in fcc metals have the lowest energy among all surfaces. For example, the surface energies of Cu {001} and {111} planes are $1.35\,\mathrm{J\,m^{-2}}$ and $1.24\,\mathrm{J\,m^{-2}}$, respectively. This difference in surface energies causes the ⟨110⟩/{111} configuration to have a lower energy and to be more stable compared with the ⟨001⟩/{001} configuration. Therefore, the ⟨001⟩/{001} wire has a natural tendency for spontaneous reorientation back to the ⟨110⟩/{111} configuration upon unloading. The reorientation essentially lowers the surface energy as a result of the increase in atomic density on surfaces when {001} surfaces reorganize into closely-packed {111} surfaces. By separating the contributions of surfaces and interior to the total strain energy of nanowires, the following analysis demonstrates the role of surfaces in driving the lattice reorientation.

Consider a wire in the ⟨001⟩/{001} configuration. The strain energy can be written as [39, 40]

$$U = U_{\mathrm{bulk}} + U_{\mathrm{surface}}, \tag{7}$$

where U_{bulk} is the strain energy in the bulk of the wire, and U_{surface} is the surface free energy. In particular, the strain energy of a square ⟨001⟩/{001} wire with lateral size d and length l (see figure 2b) is

$$U_{001} = u_{001}(\varepsilon) l d^2 + 4\gamma_{001}(\varepsilon) l d, \tag{8}$$

where $u_{001}(\varepsilon)$ is the bulk strain energy density of the wire stretched in the ⟨001⟩ orientation and $\gamma_{001}(\varepsilon)$ is the surface energy density of {001} surfaces. Figure 12a shows the total strain energy U and its two components (U_{bulk} and U_{surface}) for the ⟨001⟩/{001} wire with a cross section of $1.45 \times 1.45\,\mathrm{nm}$. The corresponding profiles for the ⟨001⟩/{001} wire with a cross-section of $1.92 \times 1.92\,\mathrm{nm}$ are shown in figure 12b. Note that these two wire configurations transform into each during loading and unloading. Clearly, U_{surface} dominates the total strain energy due to the extremely high surface-to-volume ratio of the nanowire. The self-equilibrium state of the wire without external loading can be obtained by minimizing the strain energy in equation (7) via

$$\left.\frac{\partial U_{001}}{\partial \varepsilon}\right|_{\varepsilon=\varepsilon^*} = 0, \tag{9}$$

where ε^* is the strain at the self-equilibrium state. Substitution of equation (8) into equation (9) gives

$$\left.\frac{\partial u_{001}}{\partial \varepsilon}\right|_{\varepsilon=\varepsilon^*} + \frac{4}{d}\left.\frac{\partial \gamma_{001}}{\partial \varepsilon}\right|_{\varepsilon=\varepsilon^*} = 0, \tag{10}$$

where $\sigma = \partial u_{001}/\partial \varepsilon$ is the stress in the core of the wire, and $f = \partial \gamma_{001}/\partial \varepsilon$ is the surface stress on {001} surfaces. Equation (10) can be rewritten as

$$\sigma = -\frac{4f}{d}. \tag{11}$$

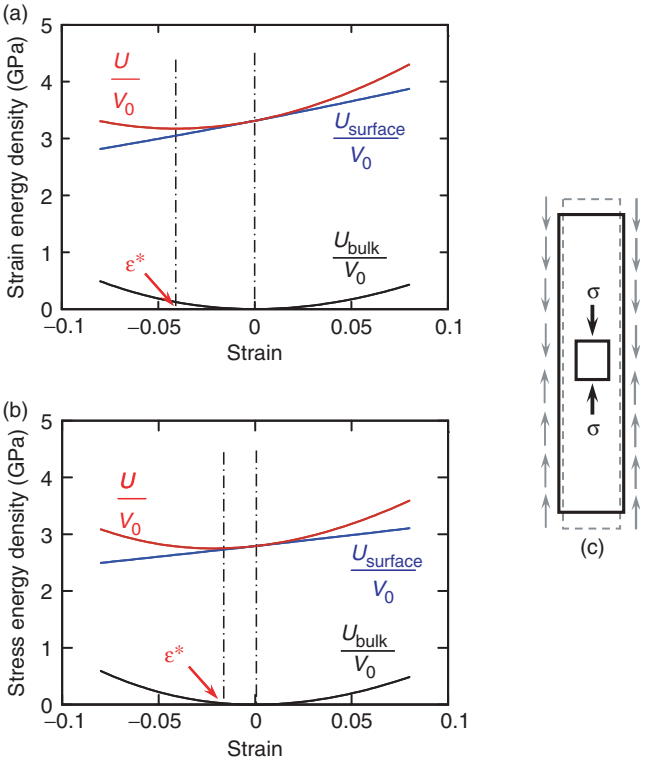

Figure 12. (a) The strain energy density function and its bulk and surface components of a 1.45×1.45 nm $\langle 001 \rangle / \{001\}$ Cu wire, the reference state is the undeformed bulk lattice; (b) the strain energy density function and its bulk and surface components of a 1.92×1.92 nm $\langle 110 \rangle / \{111\}$ Cu wire, the reference state is the undeformed bulk lattice; (c) a schematic illustration of the contraction of a wire under surface stress.

Equation (11) quantifies the compressive stress in the core of the wire at its free standing equilibrium state. This compressive stress is induced and balanced by tensile surface stress, f [39], as illustrated in figure 12c. The magnitude of σ is inversely proportional to the lateral size d of the wire. When d is very small, σ can be large enough to induce the reverse lattice reorientation which ultimately leads the wire to the more stable $\langle 110 \rangle / \{111\}$ configuration. σ can be regarded as the driving force for the spontaneous reverse lattice reorientation.

Since the $\langle 110 \rangle / \{111\}$ state always has a lower energy than the corresponding $\langle 001 \rangle / \{001\}$ state regardless of size, why does the reverse reorientation only occur at temperatures above T_{cr}? This temperature-dependence has to do with the energetic barrier and available driving force for the process. To initiate the reorientation, partial dislocations must nucleate and propagate to accommodate mobile twin boundaries. These defects are of higher energies and thus introduce an energy barrier for the initiation of reorientation process. Note that this barrier is closely related to the unstable stacking fault energy γ_{us} discussed earlier. Thermal energy can provide the necessary energy for overcoming the barrier. As wire size increases,

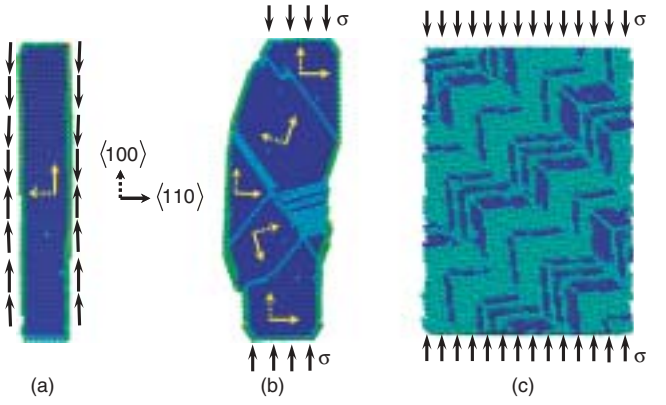

Figure 13. Effect of size on the behaviour of fcc metals: (a) a 1.45×1.45 nm ⟨001⟩/{001} Cu wire showing full reorientation and contraction under surface stress; (b) a 3.6×3.6 nm ⟨001⟩/{001} Cu wire showing partial reorientation under external compression, the yellow coordinate axes indicate the different orientations of the grains separated by twin boundaries, atoms are coloured according to their centrosymmetric values; (c) deformed configuration of a bulk Cu crystal under external compression showing 3D interlocking twin planes.

σ decreases and higher temperatures are needed to initiate the spontaneous reverse reorientation. This size- and temperature-dependence is frequently observed in experiments. For example, at a given temperature ⟨001⟩/{001} Au nanowires and nanofilms are observed to reorient into the ⟨110⟩/{111} state only when their size is reduced to less than 2 nm by electron beam irradiation [40, 41]. On the other hand, the energy barrier in Ni is much higher, therefore, the surface-stress-induced compressive stress alone may not be sufficient to initiate the spontaneous lattice reorientation even at very high temperatures. Under such conditions, externally applied compressive stress may provide the additional stimulus needed to overcome the barrier.

5.2. *Size effect on shape memory behaviour*

Bulk fcc metals have not been known to possess shape memory. One logical question is why fcc nanowires exhibit shape memory and pseudoelasticity while their bulk counterparts do not. The reason lies in their extremely small sizes at the nanometre scale and their unique 1D structure. First of all, it has been demonstrated that twinning is vital to the pseudoelasticity in the nanowires. Slip is generally favoured over twinning in bulk fcc metals under normal loading conditions. Twinning becomes a viable deformation mode only when the dominant size (e.g. grain size or wire size) scale approaches nanometres. Specifically, both experiments and simulations have shown that twinning is the primary deformation mechanism in nanocrystalline metals, nanowires, and nanoparticles [37]. Secondly, nanowires have extremely large surface-to-volume ratios. For example, a nanowire has a surface-to-volume ratio 10^6 times that of a macroscopic specimen [16]. The fraction of surface atoms is around 39.5% for a 1.8×1.8 nm nanowire, whereas that for bulk metals

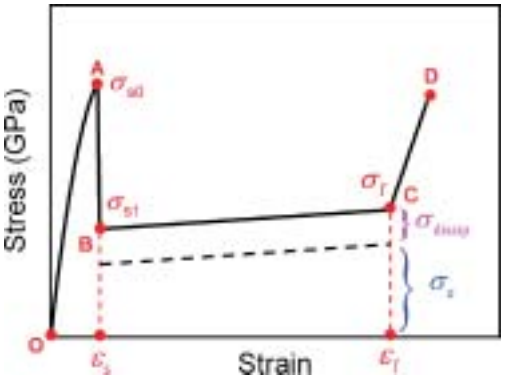

Figure 14. A schematic illustration of the stress–strain curve of a nanowire under isothermal quasistatic tensile loading.

approaches zero. The high surface-to-volume ratios induce high values of internal compressive stress σ which are on the order of GPa, providing the necessary driving force for the reverse reorientation at temperatures above T_{cr}, even in the absence of externally applied forces, as shown in figure 13a. On the other hand, σ is only on the order of Pascals in bulk materials, far from being sufficient to initiate spontaneous lattice reorientations. One may still wonder if a bulk fcc crystal with the $\langle 110 \rangle / \{111\}$ configuration would exhibit the lattice reorientation leading to pseudoelasticity and shape memory in nanowires under the condition that an external compressive stress of sufficient magnitude is applied. The answer is clearly "no" because the reversibility of the reorientation process primarily results from the unique 1D structure at the nanoscale. This structure acts as a channel, limiting the propagation of twin boundaries to the direction of the wire axis. The opposite directions of twin boundary propagation during loading and unloading give rise to the reversible nature of deformations, as shown in figure 13a. As the lateral dimensions increase, the wire gradually becomes a 3D structure with multiple modes of defect nucleation, propagation and interaction. Such a 3D structure leads to simultaneous twin boundaries and stacking faults at multiple sites and in multiple directions, as shown in figures 13b and c. The entanglement and interactions of these defects reduce the mobility of the twin boundaries and ultimately preclude the occurrence of the pseudoelasticity and the shape memory. Under such conditions, the wires are essentially a polycrystal with grains separated by interlocking twin boundaries.

6. Micromechanical continuum model

We now turn the discussion to the development of a micromechanical continuum model for characterizing the novel tensile behaviour of the shape memory nanowires. This model is piecewise smooth with each segment associated with one of the deformation stages discussed previously in section 2. The model captures the major

characteristics of the lattice reorientation process, especially all the critical parameters (see figure 14), including

(1) the yield stress σ_{s0} at which the initiation of lattice reorientation occurs;
(2) the start and finish strains (ε_s and ε_f) for the lattice reorientation process; and
(3) the stress levels throughout lattice reorientation process, especially the start and finish stresses (σ_{s1} in figure 14 and σ_f).

Since the two wire configurations involved have different strain energies at the unstressed state, they are also referred to as phases here. For clarity, it is important to define the stress measures and strain measures involved in this analysis. In particular, the stress σ and strain ε associated with the stress-strain curve in figure 4 are the average stress and average strain within the whole wire, although local stresses and strains in the two phases are not necessarily the same. More precisely, ε is the nominal engineering strain with the initial unstressed $\langle 110\rangle/\{111\}$ wire as its reference state, i.e.

$$\varepsilon = \frac{l - l_0}{l_0}. \tag{12}$$

Here, l_0 is the length of the initial, unstressed $\langle 110\rangle/\{111\}$ wire at a given temperature, l is the current length of the stretched wire. The total mechanical work done by external load is

$$W = V_0 \int_0^{\varepsilon} \sigma d\varepsilon. \tag{13}$$

Accordingly, the average stress σ is

$$\sigma = \frac{1}{V_0}\frac{\partial W}{\partial \varepsilon}, \tag{14}$$

where V_0 is the initial wire volume at the unstressed state. Since stresses and strains are not necessarily uniform throughout the transforming wire, the local stresses (σ_{110} and σ_{001}) and strains (ε_{110} and ε_{001}) in the $\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ phases, respectively, are considered, as shown in figure 5a. Here, ε_{110} and ε_{001} are engineering strains with the reference states being the corresponding unstressed self-equilibrium states of the $\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ phases, respectively.

The following discussions focus on the lattice reorientation process because it is this unique deformation mechanism that leads to the novel SME of metal wires with large reversible strains. For the two elastic stages of deformation (OA and CD in figures 4 and 14), the stress–strain relation satisfies the constitutive equations of the pure $\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ phases, respectively, i.e.

$$\sigma_{110} = \frac{du_{110}}{d\varepsilon_{110}}, \tag{15}$$

and

$$\sigma_{001} = \frac{du_{001}}{d\varepsilon_{001}}. \tag{16}$$

In the above expressions, u_{110} and u_{001} are the strain energy density functions for the two phases.

6.1. *Decomposition of the lattice reorientation process*

The lattice reorientation process involves both strain energy transfer between the two phases and energy dissipation, even at infinitesimally low deformation rates (i.e. under quasistatic conditions). To reveal the mechanisms, this process is decomposed into a non-dissipative part and a dissipative part. The non-dissipative part concerns the smooth transition between a series of phase-equilibrium states. This part is an ideal and thermodynamically reversible process. The corresponding portion of the input mechanical work is stored as strain energy in the two phases of the wire and can flow between the phases as one phase transforms into the other. The dissipative part concerns the process of twin boundary propagation (dislocation nucleation, propagation, and annihilation) which involves overcoming the energy barriers between phase-equilibrium states. This part is a thermodynamically irreversible process. For the reversible part, we assume the wire goes through the series of phase-equilibrium states *smoothly*, although the actual transitions between equilibrium states are not smooth because barriers exist for lattice-scale defect nucleation. The wire is periodically brought to unstable states at the imminence of the nucleation of partial dislocations on a side surface. Following the dislocation nucleation and annihilation on the other free surface, the wire settles into the phase-equilibrium states and remains in such smooth-changing states between the nucleation and annihilation events. The part of the input work that is associated with the dislocation nucleation, propagation, and annihilation is dissipated. The first law of thermodynamics dictates that

$$W = \Delta U + Q, \tag{17}$$

where W is the total input work, Q is the energy dissipation associated with the irreversible process. ΔU is the change in strain energy associated with the reversible process, i.e.,

$$\Delta U = U - U_0 \tag{18}$$

with U being the strain energy associated with the phase-equilibrium state and U_0 being the initial strain energy of the unstressed $\langle 110\rangle/\{111\}$ wire. Equations (14) and (17) combine to give

$$\sigma = \frac{1}{V_0}\left(\frac{\partial(\Delta U)}{\partial \varepsilon} + \frac{\partial Q}{\partial \varepsilon}\right) = \sigma_{\mathrm{e}} + \sigma_{\mathrm{dissip}}, \tag{19}$$

where σ_{e} is the part of the stress needed to drive the transition of the phase-equilibrium states. It is determined by the elastic properties of the two phases. Specifically,

$$\sigma_{\mathrm{e}} = \frac{1}{V_0}\frac{\partial(\Delta U)}{\partial \varepsilon} = \frac{1}{V_0}\frac{\partial(U - U_0)}{\partial \varepsilon} = \frac{1}{V_0}\frac{\partial U}{\partial \varepsilon}. \tag{20}$$

σ_{dissip} is the part of the stress required to drive or absorb the nucleation, propagation, and annihilation of partial dislocations associated with the twin boundary propagation. Obviously,

$$\sigma_{\text{dissip}} = \frac{1}{V_0}\frac{\partial Q}{\partial \varepsilon}. \tag{21}$$

The above decomposition of energy and stress allows the conservative and dissipative parts of the deformation process to be analyzed separately. To carry out this analysis, the governing equations for the phase-equilibrium states of the wire are first derived based on a constrained minimization of the strain energy for the wire. The reversible smooth transition process is obtained by solving these governing equations for a series of phase-equilibrium states between the onset and completion. The source and mechanism of the energy dissipation process are then analyzed from the perspective of lattice-scale dislocation activities and the characteristics of the associated energy landscape, allowing the dissipative process to be quantified.

6.2. *Conservative smooth transition between phase-equilibrium states*

As the $\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ configurations correspond to two energy wells of the strain energy function for the transforming nanowire, the non-dissipative part is modelled within the framework of strain energy functions with multiple local minima of [42–45]. Within this framework, at any given transformation state quantified by the average strain ε the wire adopts the two-phase wire structure that has the minimum strain energy, under the condition that the kinematical constraints and force equilibrium are satisfied.

Consider such a phase-equilibrium state with nominal strain ε and two coexisting phases separated by a twin boundary, as shown in figure 5a. Each phase is elastically stretched relative to its own free-standing equilibrium (reference) state. The following field equations must be satisfied.

$$\varepsilon_i = \frac{d\delta_i}{dx} \text{ in each phase,} \tag{22}$$

and

$$\delta^+ = \delta^- \text{ at the twin boundary.} \tag{23}$$

Here, δ_i is the axial displacement of the wire. For the $\langle 110\rangle/\{111\}$ phase, $i = 110$ and $\delta_i = \delta_{110}$; for the $\langle 001\rangle/\{001\}$ phase, $i = 110$ and $\delta_i = \delta_{001}$. Similarly, ε_i is the local strain in the wire with $\varepsilon_i = \varepsilon_{110}$ in the $\langle 110\rangle/\{111\}$ phase and $\varepsilon_i = \varepsilon_{001}$ in the $\langle 001\rangle/\{001\}$ phase. δ^+ and δ^- denote the limiting values of the displacement at the twin interface with the limits being taken towards the interface from either side. Equations (22) and (23) require that the displacement be continuous within each phase and across the twin interface. Although the strain is continuous within each phase, it is different in the two phases in general and a strain discontinuity exists across the twin interface.

It is important to point out that the total strain ε includes not only contributions from the elastic strains in the two phases (ε_{110} and ε_{001}), but also the transformation strain ($\varepsilon_{\mathrm{tr}}$) due to the phase transformation. Specifically, note that the current wire length is

$$l = l^0_{110}(1+\varepsilon_{110}) + l^0_{001}(1+\varepsilon_{001}), \tag{24}$$

where l^0_{110} and l^0_{001} and are the lengths of the $\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ phases at their unstressed states, respectively; $l^0_{110}(1+\varepsilon_{110})$ and $l^0_{001}(1+\varepsilon_{001})$ are the current lengths of the $\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ phases at a nominal wire strain of ε, respectively. Moreover, the sum of the lengths of the untransformed $\langle 110\rangle/\{111\}$ phase and the transformed $\langle 110\rangle/\{111\}$ phase must be equal to the total length of the initial undeformed wire, i.e.

$$l^0_{110} + \frac{l^0_{001}}{1+\varepsilon_{\mathrm{tr}}} = l_0, \tag{25}$$

where $l^0_{001}/(1+\varepsilon_{\mathrm{tr}})$ is the length of the transformed $\langle 110\rangle/\{111\}$ phase calculated from the length of the corresponding $\langle 001\rangle/\{001\}$ phase.

In addition to the kinematical constraints, the stress field in each phase must satisfy the force balance condition, i.e.

$$\frac{d\sigma_i}{dx} = 0 \text{ in each phase}, \tag{26}$$

and

$$\sigma_{110}A_{110} = \sigma_{001}A_{001} \text{ at the phase boundary}, \tag{27}$$

where σ_i is the axial stress in the wire. For the $\langle 110\rangle/\{111\}$ phase, $\sigma_i = \sigma_{110}$; for the $\langle 001\rangle/\{001\}$ phase, $\sigma_i = \sigma_{001}$. A_{110} is the area of the rhombic cross-section of the $\langle 110\rangle/\{111\}$ phase, and A_{001} is the area of the square cross-section of the $\langle 001\rangle/\{001\}$ phase. Both A_{110} and A_{001} are cross-sectional areas in the unstressed states. Obviously, the axial force is constant within each phase and across the twin interface. The stress is uniform within each phase but discontinuous across the twin interface because the cross-sectional areas of the two phases are not the same ($A_{110} > A_{001}$).

The total strain energy of the wire is composed of the strain energy of each phase and the energy of the interface, i.e.

$$U = \int_{V_{110}} u_{110}(\varepsilon_{110})dV + \int_{V_{001}} u_{001}(\varepsilon_{001})dV + U_{\mathrm{interface}}. \tag{28}$$

The local stresses, strains, strain energies and lengths of the two phases at a given strain ε can be obtained by minimizing U under the kinematical and kinetic constraints of equations (22)–(28). The complete smooth transition process can be characterized by repeating the solution process for the strain range of $\varepsilon_{\mathrm{s}} \le \varepsilon \le \varepsilon_{\mathrm{f}}$ (figure 14). Note that the strain energy associated with the twin interface $U_{\mathrm{interface}}$ is essentially constant because the structure of the coherent boundary remains the same as it propagates through the wire. Consequently, the specific value of $U_{\mathrm{interface}}$ does

not affect the energy minimization process and the predicted values of the stresses, strains, and sizes of the phases.

The part of the stress needed to drive the transition of the phase-equilibrium states σ_e can be obtained by substituting equation (28) into equation (20). The result is

$$\sigma_e = \lambda_{110}\sigma_{110} + \lambda_{001}\sigma_{001}, \tag{29}$$

where

$$\lambda_{110} = \frac{l^0_{110}}{l_0}\frac{l^0_{110} + l^0_{001}}{l^0_{110} + (A_{110}/A_{001})((d\sigma_{110}/d\varepsilon_{110})/(d\sigma_{001}/d\varepsilon_{001}))l^0_{001}} \tag{30}$$

and

$$\lambda_{001} = \frac{l^0_{001}}{l_0}\frac{l^0_{110} + l^0_{001}}{((d\sigma_{001}/d\varepsilon_{001})/(d\sigma_{110}/d\varepsilon_{110}))l^0_{110} + (A_{110}/A_{001})l^0_{001}}. \tag{31}$$

At the initiation of the lattice reorientation process (point B in figure 14),

$$\sigma_e = \sigma_{110}. \tag{32}$$

Similarly, at the completion of the lattice reorientation process (point C in figure 14),

$$\sigma_e = \sigma_f = (1 + \varepsilon_{tr})\left(\frac{A_{001}}{A_{110}}\right)\sigma_{001}. \tag{33}$$

Since $A_{110} > A_{001}$, equation (27) shows that $\sigma_{110} < \sigma_{001}$. Therefore, σ_e takes on its smallest value σ_{110} at point B in figure 14 and its maximum value $(1 + \varepsilon_{tr})(A_{001}/A_{110})\sigma_{001}$ at point C in figure 14. As the transformation progresses from B to C, σ_e increases as the lower stress $\langle 110\rangle$ phase is gradually converted into the higher stress $\langle 001\rangle$ phase. This increase of σ_e is the primary reason behind the increase of $\sigma = \sigma_e + \sigma_{dissip}$ over the same stage of deformation.

6.3. *Dissipative twin boundary propagation*

The smooth transition between phase-equilibrium states discussed in the preceding subsection is an idealized process which is thermodynamically reversible without energy dissipation. This part of the model focuses on the transformation outcome and its energy requirement and does not account for the transformation path or how the transformation occurs. To state this differently, the analysis so far has not considered the effects of the energy barrier between the$\langle 110\rangle/\{111\}$ and $\langle 001\rangle/\{001\}$ states. The actual transformation is a dissipative process regardless of how slowly the transformation occurs. In general, the dissipation during phase transformations is due to the ruggedness of the energy landscape associated with barriers between energy wells corresponding to different phases. These barriers exist because energy is required to nucleate defects or structural changes which are necessary for the transformation to occur [46–49]. For the fcc nanowires, the energy required to form $(1/6)\langle 112\rangle$Shockley partial dislocations associated with the twin boundary constitutes the primarily part of the energy barrier. The propagation of the twin

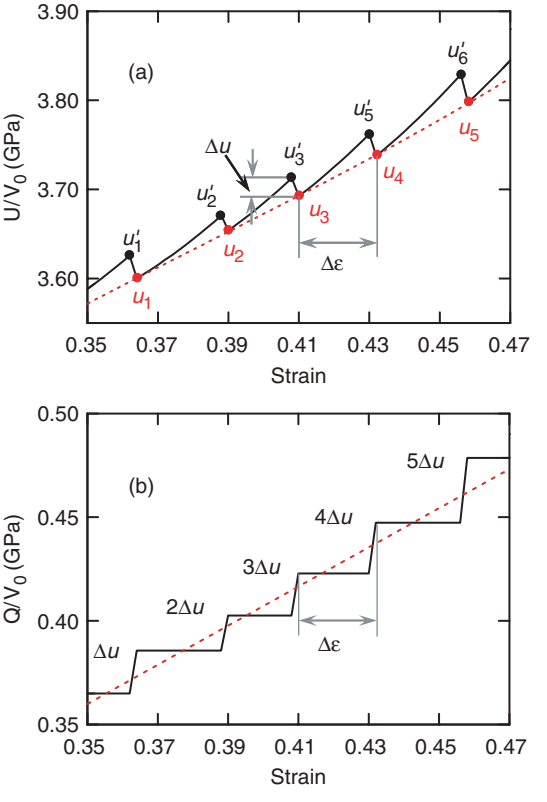

Figure 15. (a) The serrated strain energy curve during lattice reorientation, the local peaks correspond to unstable states created upon the nucleation of partial a dislocation, and the local minima correspond to phase equilibrium states; (b) the stepwise energy dissipation curve.

boundary along the wire axis through sequential nucleation, gliding, and annihilation of the partial dislocations [45] assigns the strain energy function profile a serrated shape with local peaks (u_i') corresponding to unstable states and local minima (u_i) corresponding to metastable phase-equilibrium states, as shown in figure 15. The twin boundary propagation is in a "stick-slip" manner [48]. At the same time, the wire is periodically brought into unstable states and subsequently settles into metastable states after overcoming the barrier. The local minima u_i correspond to the metastable states and local maxima u_i' correspond to the unstable states. It is important to point out that the dotted line in figure 15a illustrates the reversible smooth transition path between the metastable states governed by the equations in subsection 6.2. On the other hand, the serrated strain energy density curve shows the actual change of strain energy density. Each time when a dislocation is nucleated and glides across the wire, the strain energy drops precipitously from the local maximum u_i' to the minimum u_i. At the same time, the strain energy difference $\Delta u = u_i' - u_i$ is dissipated out of the wire through heat transfer, acoustic wave propagation, and irradiation. If thermal fluctuations are neglected, Δu is essentially constant because the energy barrier for dislocation nucleation is the same for all

"stick-slip" steps throughout the transformation. When the twin boundary is in the "stick" mode (not propagating), the wire undergoes elastic stretching and all input energy is stored as strain energy. Therefore, no energy is dissipated. This process causes Q to increase in a stepwise manner as a function of tensile strain ε, as shown in figure 15b. This stepwise increasing profile can be approximated by the dotted line in figure 15b which is analytically

$$Q = V_0 \frac{n - 0.5}{n} \frac{\Delta u}{\Delta \varepsilon} \varepsilon \approx V_0 \frac{\Delta u}{\Delta \varepsilon} \varepsilon, \tag{34}$$

where n is the total number of steps in the plot; $\Delta\varepsilon$ is the strain increment corresponding to one cycle of dislocation nucleation and annihilation (or one step of twin boundary propagation) which is a constant. Substitution of equation (34) into equation (21) gives

$$\sigma_{\text{dissip}} = \frac{\Delta u}{\Delta \varepsilon}. \tag{35}$$

This is a constant that should be regarded as the average value of the stress required to overcome the energy barriers and nucleate a dislocation. Note that σ_{dissip} is proportional to Δu which is closely related to the energy barrier for the nucleation of dislocations associated with the propagation of the twin interface.

As thermal fluctuations facilitate the overcoming of the energy barriers, at higher temperatures less mechanical work or lower stresses are needed for the nucleation of dislocations. Therefore, Δu and σ_{dissip} are smaller at higher temperatures.

6.4. *Comparison to the results of MD calculations*

The model predictions are compared with the results of MD simulations for a Cu nanowire at three different temperatures between 100 and 300 K. In addition, the results for wires with three different sizes between 1.5 and 2.4 nm are also compared. The MD calculations are carried out with the EAM interatomic potential for Cu by Mishin *et al.* [50, 51] under isothermal, quasistatic conditions. Further details of the calculations can be found elsewhere [6].

The transformation process is temperature-dependent, with both the yield stress σ_{s0} and the transformation stress σ decreasing as temperature increases. Figure 16 shows the stress–strain curves at 100, 200 and 300 K for a $\langle 110 \rangle / \{111\}$ Cu nanowire with a cross-section of 1.96×1.96 nm. The model predictions are in excellent agreement with the results of MD calculations at all temperatures shown. Specifically, the effect of temperature is primarily on σ_{dissip} which decreases as temperature increases. Obviously, this trend has to do with the fact that thermal fluctuations facilitate the overcoming of the energy barriers for dislocation nucleation. For the same reason, the yield stress is lower at higher temperatures. In contrast, the part of the total stress associated with the conservative process of strain energy reverberation between the phases (σ_e) is essentially independent of temperature. This is because σ_e is primarily determined by the elastic properties of the two phases which are essentially identical at the different temperatures analyzed.

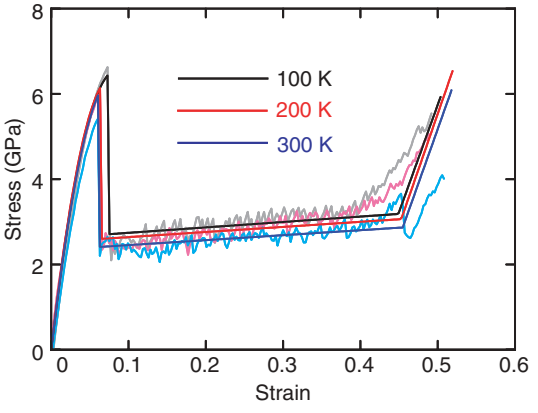

Figure 16. Comparison of the stress–strain curves of a 1.96×1.96 nm $\langle 110 \rangle / \{111\}$ Cu wire at 100, 200 and 300 K.

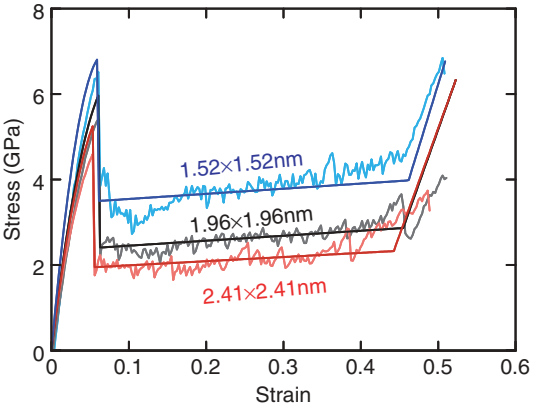

Figure 17. Comparison of stress–strain curves of $\langle 110 \rangle / \{111\}$ Cu wires with three different sizes at 300 K.

The model captures quite accurately the significant size effects observed in the tensile behaviour of $\langle 110 \rangle / \{111\}$ wires as well. Figure 17 shows the stress–strain curves of three wires with cross-sectional sizes of 1.52×1.52 nm, 1.96×1.96 nm and 2.41×2.41 nm at 300 K. In particular, note that the transformation stress σ decreases as wire size increases. The model predictions show excellent agreement with the results of MD simulations. This accuracy of the model allows it to help explain the size dependence of the tensile behaviour. Specifically, smaller wires are stiffer than larger wires due to the higher surface-to-volume ratios at smaller sizes. Consequently, σ_e is higher in smaller wires because it is primarily determined by the elastic properties of the two phases. On the other hand, σ_{dissip} is proportional to $\Delta u = \Delta U / V_0$ [equation (35)], where ΔU is the unit energy dissipation induced by the nucleation and annihilation of a single dislocation. Since this unit dissipation is essentially independent of the size of the wires, σ_{dissip} is inversely proportional to the

wire volume. The fact that both σ_e and σ_{dissip} decrease as wire size increases explains why the overall transformation stress σ is lower at larger wire sizes.

7. Conclusions

A novel SME and its underlying pseudoelastic behaviour have been discovered recently in single-crystalline fcc metal (Cu, Ni and Ag) nanowires. These wires can recover axial elongations of up to 50% which are well beyond the recoverable strains of 5–8% typical for most bulk SMAs. Results of atomistic simulations and evidences from experiments show that this phenomenon only exists at the nanometre scale and is associated with a reversible lattice reorientation driven by the high surface-stress-induced internal stresses at the nanoscale. This novel SME exists over a wide range of temperature and is associated with response times on the order of nanoseconds, making the nanowires attractive functional components for a new generation of biosensors, transducers, and interconnects in NEMS.

It is understood that not all fcc metal wires show this novel shape memory. Only wires with lateral sizes of up to approximately 5 nm made of fcc metals with higher tendencies for twinning (such as Cu and Ni) possess this shape memory. On the other hand, fcc metals with lower propensities for twinning (such as Al) do not possess this shape memory. It is possible for wires of some fcc metals with intermediate propensities for twinning (e.g. Au) to show a transition of behaviour from pseudoelasticity at lower temperatures to plasticity at higher temperatures. The differences in behaviour among wires of different materials are due to the two competing deformation mechanisms of twinning and slip. Specifically, the pseudoelastic behaviour of Cu and Ni wires is associated with a reversible lattice reorientation through twinning. On the other hand, the irreversible tensile deformation of Al wires is due to crystalline slip. The transitional behaviour of Au wires is due to a change in deformation mechanism from twinning at lower temperatures to slip at higher temperatures.

The novel shape memory behaviour only exists at the nanoscale but not in bulk metals. This explains why it has not been discovered before because extensive research on nanowires has been carried out only in the last decade or so. For this behaviour to be observed, a combination of three conditions must exist. These conditions are: (i) the characteristic size scale must be comparable to the width of the stable stacking faults of the FCC metals so that twinning can become the favoured deformation mechanism – this size scale is between a few nanometres to tens of nanometres; (ii) surface-stress-induced internal compressive stress must be high enough to cause the initiation of a spontaneous lattice reorientation without external influence; this, again, introduces a size scale on the order of a few nanometres; and (iii) properly oriented 1D or quasi-1D structure must be available to channel unimpeded propagation of a small number of untangled twin boundaries. Obviously, nanowires of Cu, Ni, and Au with lateral dimensions up to 5 nm and the $\langle 110\rangle/\{111\}$ configuration meet the above criteria at appropriate temperatures. It remains to be seen if wires of other fcc metals (e.g. Ag and Pt) possess the pseudoelasticity and SME as well.

A micromechanical continuum model has been developed for the unique pseudoelastic behaviour of the nanowires during quasistatic tensile loading. In this model, the lattice reorientation process is decomposed into a reversible smooth transition part between phase-equilibrium states and a dissipative twin boundary propagation part. The smooth transition part between metastable phase-equilibrium states is modelled using strain energy functions with multiple local minima. These functions are obtained directly from the interatomic potentials and relevant wire structures. The dissipative nature of the twin boundary propagation process is due to the ruggedness of the strain energy landscapes associated with the nucleation, gliding, and annihilation of partial dislocations. Specifically, each time a dislocation is nucleated and annihilated, the strain energy difference between the unstable state upon nucleation and the equilibrium state after annihilation is dissipated out of the nanowire in the form of heat. This model captures the major characteristics of the novel pseudoelastic behaviour and accurately accounts for related size- and temperature-dependence of the behaviour.

Acknowledgements

This research is supported by NASA Langley Research Center through grant # NAG-1-02054. Computations are carried out at the NAVO, ERDC and ARL MSRCs through AFOSR MURI # D49620-02-1-0382. We thank S. Plimpton for sharing his MD code WARP [52]. Images of deformation in this paper are created with the graphics package visual molecular dynamics (VMD) [53].

References

[1] K. Otsuka and C.M. Wayman, *Shape Memory Materials* (Cambridge University Press, New York, 1998), p. 282.
[2] W.W. Liang and M. Zhou, J. Engng. Mater. Technol. **127** 423 (2005).
[3] W.W. Liang, M. Zhou and F.J. Ke, Nano. Lett. **5** 2039 (2005).
[4] H.S. Park, K. Gall and J.A. Zimmerman, Phys. Rev. Lett. **95** 255504 (2005).
[5] H.S. Park and C. Ji, Acta. Mater. **54** 2645 (2006).
[6] W. Liang and M. Zhou, Phys. Rev. B **73** 115409 (2006).
[7] E.B. Tadmor and S. Hai, J. Mech. Phys. Solids **51** 765 (2003).
[8] H.S. Park, K. Gall and J.A. Zimmerman, J. Mech. Phys. Solids **54** 1862 (2006).
[9] Z.L. Wang, R.P. Gao, B. Nikoobakht, *et al*., J. Phys. Chem. B **104** 5417 (2000).
[10] V. Rodrigues, J. Bettini, A.R. Rocha, *et al*., Phys. Rev. B **65** 153402 (2002).
[11] Z. Liu and Y. Bando, Adv. Mater. **15** 303 (2003).
[12] Z. Liu, Y. Yang, J. Liang, *et al*., J. Phys. Chem. B **107** 12658 (2003).
[13] Y. Wang, J. Yang, C. Ye, *et al*., Nanotechnology **15** 1437 (2004).
[14] J. Diao, K. Gall and M.L. Dunn, Phys. Rev. B **70** 075413 (2004).
[15] J. Wang, H. Huang and T.S. Cale, Modell. Simul. Mater. Sci. Engng **12** 1209 (2004).
[16] W. Liang and M. Zhou, J. Engng. Mater. Technol. **127** 423 (2005).
[17] H.S. Park and J.A. Zimmerman, Scripta Mater. **54** 1127 (2006).

[18] V. Rodrigues and D. Ugarte, Nanotechnology **13** 404 (2002).
[19] J.C. González, V. Rodrigues, J. Bettini, *et al*., Phys. Rev. Lett. **93** 126103-1 (2004).
[20] Y. Kondo and K. Takayanagi, Phys. Rev. Lett. **79** 3455 (1997).
[21] W. Liang and M. Zhou, J. Mech. Engng. Sci. **218** 599 (2004).
[22] W. Liang, M. Zhou and F. Ke, Nano. Lett. **5** 2039 (2005).
[23] S. Nosé, Molec. Phys. **52** 255 (1984).
[24] W.G. Hoover, Phys. Rev. A **31** 1695 (1985).
[25] K. Gall, J. Diao and M.L. Dunn., Nano. Lett. **4** 2431 (2004).
[26] C.B. Carter and I.L.F. Ray, Phil. Mag. **35** 189 (1977).
[27] R.W. Balluffi, J. Nucl. Mater. **69–70** 240 (1978).
[28] W.M. Stobbs and C.H. Sworn, Phil. Mag. **24** 1365 (1971).
[29] L.E. Murr, *Interfacial Phenomena in Metals and Alloys* (Addison-Wesley, Reading, MA, 1975).
[30] K. Westmaco and R.L. Peck, Phil. Mag. **23** 611 (1971).
[31] R.H. Rautioaho, Phys. Stat. Sol. (b) **112** 83 (1982).
[32] H.V. Swygenhoven, P.M. Derlet and A.G. Frøseth, Nature Mater. **3** 399 (2004).
[33] J.R. Rice, J. Mech. Phys. Solids **40** 239 (1992).
[34] E.B. Tadmor and N. Bernstein, J. Mech. Phys. Solids **52** 2507 (2004).
[35] R.W. Hertzberg, *Deformation and Fracture Mechanics of Engineering Materials,* 3rd edn (Wiley, New York, 1989).
[36] C. Meyers, *Mechanical Metallurgy Principle and Applications* (Printice-Hall, Englewood Cliffs, NJ, 1984).
[37] D.P. Field, B.W. True, T.M. Lillo, *et al*., Mater. Sci. Engng. A **372** 173 (2004).
[38] J.P. Hirth and J. Lothe, *Theory of Dislocations* (John Wiley & Sons, New York, 1982), p. 315.
[39] F.H. Streitz, R.C. Cammarata and K. Sieradzki, Phys. Rev. B **49** 10699 (1994).
[40] R. Dingreville, J. Qu and M. Cherkaoui, J. Mech. Phys. Solids **53** 1827 (2005).
[41] Y. Kondo, Q. Ru and K. Takayanagi, Phys. Rev. Lett. **82** 751 (1999).
[42] R. Abeyaratne and J.K. Knowles, J. Mech. Phys. Solids **41** 541 (1993).
[43] R. Abeyaratne and S.-J. Kim, Int. J. Solids Struct. **31** 2229 (1994).
[44] R. Abeyaratne, K. Bhattacharya and J.K. Knowles, *Nonlinear Elasticity: Theory and Applications* (Cambridge University Press, Cambridge, 2001).
[45] R. Abeyaratne and S. Vedantam, J. Mech. Phys. Solids **51** 1675 (2003).
[46] L. Truskinovsky and A. Vainchtein, Phys. Rev. B **67** 172103-1-4 (2003).
[47] L. Truskinovsky and A. Vainchtein, J. Mech. Phys. Solids **52** 1421 (2004).
[48] A. Vainchtein and P. Rosakis, J. Nonlinear Sci. **9** 697 (1999).
[49] G. Puglisi and L. Truskinovsky, Continuum. Mech. Thermodynamics **14** 437 (2002).
[50] Y. Mishin, D. Farkas, M.J. Mehl, *et al*., Phys. Rev. B **59** 3393 (1999).
[51] Y. Mishin, M.J. Mehl, D.A. Papaconstantopoulos, *et al*., Phys. Rev. B **63** 224106 (2001).
[52] S.J. Plimpton, J. Comput. Phys. **117** 1 (1995).
[53] W. Humphrey, A. Dalke and K. Schulten, J. Molec. Graphs **14** 33 (1996).